

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

AC

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :  C12N 15/70		A1	(11) International Publication Number: WO 00/66756  (43) International Publication Date: 9 November 2000 (09.11.00)			
(21) International Application Number: PCT/FI00/00387			(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).			
(22) International Filing Date: 3 May 2000 (03.05.00)						
(30) Priority Data: 991014 4 May 1999 (04.05.99) FI						
(71)(72) Applicants and Inventors: KORPELA, Timo [FI/FI]; Kasarminkatu 5-8, FIN-20500 Turku (FI). MACINTYRE-AYANE, Sheila [GB/GB]; 10 Barnsdale Road, Reading RG2 7JN (GB). ZAVIALOV, Anton Vladimirovich [RU/RU]; Profsouznaya str., 98-4-58, Moscow, 117485 (RU). BATTCHIKOVA, Natalia Vsevolodovna [RU/FI]; Tikankatu 28-C-13, FIN-20610 Turku (FI). PETROVSKAYA, Lada Evgenievna [RU/RU]; Vucheticha str., 44-1, Moscow, 125206 (RU). KOROBKO, Vyacheslav Grigorievich [RU/RU]; Tsurupa str., 7-2-28, Moscow, 117418 (RU). ZAV'YALOV, Vladimir Petrovich [RU/RU]; Zavodskaya str., 10-12, Lyubuchany, Chekhov District, Moscow regi, 142380 (RU).						
(74) Common Representative: KORPELA, Timo; University of Turku, BioCity 6th floor, FIN-20520 Turku (FI).						
(54) Title: MICROBIAL PROTEIN EXPRESSION SYSTEM						
(57) Abstract						
<p>The present invention provides bacterial strains for secretion of soluble biologically active recombinant heterologous proteins into periplasm or on a surface/into a cultivation medium of bacteria. The invention exploits the secretion system of Gram-negative bacteria including periplasmic chaperones and usher/secretin proteins. For accomplishing the aim, the bacterial strains simultaneously express the fusion protein (signal peptide)-(mature heterologous protein)-(subunit of a bacterial surface structure, Caf1), periplasmic chaperone specific for the subunit, and outer membrane usher/secretin protein specific for the subunit. Secretion of fusion proteins: (signal peptide of Caf1)-(mature human IL-1<math>\beta</math>)-(mature Caf1), (signal peptide of Caf1)-(mature human GM-CSF)-(mature Caf1), and (signal peptide of Caf1)-(mature human IL-1ra)-(mature Caf1) that were expressed in <i>Escherichia coli</i> simultaneously with the periplasmic chaperone Caf1M and the usher/secretin protein Caf1A are examples of the use of the invention.</p>						

BEST AVAILABLE COPY

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## MICROBIAL PROTEIN EXPRESSION SYSTEM

### 5 FIELD OF INVENTION

This invention is related to biotechnology and more specifically to production of recombinant heterologous proteins by microbes. In particular, this invention concerns secretion of soluble biologically active heterologous proteins into 10 periplasm and/or on a surface/into a cultivation medium of Gram-negative bacteria. The invention exploits the secretion system of Gram-negative bacteria including periplasmic chaperones and usher/secretin proteins of the system.

### BACKGROUND OF INVENTION

15 Commercial production of various medically and industrially valuable recombinant proteins by microbes is one the key challenges of modern biotechnology. Even though such systems are known, there are severe technical problems which are encountered within large-scale exploitation of microbial cell 20 machinery. There are several secretion systems in Gram-negative bacteria which can be potentially exploited for secretion of recombinant heterologous proteins. The systems are briefly reviewed below.

#### *1. Secretion across the inner membrane*

25 The majority of secreted proteins in *Escherichia coli* are synthesized as precursors with a classic N-terminal signal peptide (SP) which is essential for efficient export and which is cleaved during or following translocation across the inner membrane. Translocation is mediated by the Sec translocase (SecA/Y/G/E). SecA is a 30 peripherally associated ATPase, which interacts with the signal sequence and mature part of the precursor, guides the polypeptide into the translocator and provides energy for the process. Sec Y, E and G are integral membrane proteins

that form the translocator itself and a central aqueous channel through which the polypeptide is translocated. SecD and F have large periplasmic domains. Proposed functions of these two proteins are related to later steps in the process and have included release from the membrane and mediation of transfer of energy from the 5 proton motive force. Polypeptides are translocated in an 'unfolded' state. Hence SecB is a cytosolic chaperone apparently dedicated to the secretion pathway which is required for export of a subset of secreted proteins. SecB both inhibits premature folding and targets the precursor to the membrane translocase complex. It now appears that basic principles of the export system are universal. Although 10 translocation is primarily co-translational in eukaryotic systems and targeting to the secretion apparatus is primarily via SRP (signal recognition particle), the core translocator is homologous in both systems (yeast Sec 61  $\alpha$  and  $\gamma$  are homologues of *E. coli* SecY/E). Also, *E. coli* *ffh* and 4.5S RNA are homologues of eukaryotic 54 KD subunit and 7S RNA of SRP, respectively. Comparison of the eukaryotic 15 and prokaryotic systems has been extensively reviewed recently (Rapoport, T., et al. (1996)Annual Review of Biochemistry. 65:271-303; Schatz, G., and B. Dobberstein (1996)Science. 271:1519-1526).

Similarities in the basic function of eukaryotic and bacterial export systems have 20 meant that some mammalian proteins can be successfully secreted to the periplasm of *E. coli*. Examples include human insulin. Often, however, fine tuning is required such as optimising the N-terminus of the mature protein, removal of positively charged residues at the end of the SP or beginning of the mature protein. 25 ensuring presence of a good cleavage site. Frequently, a bacterial signal peptide such as the OmpA SP has been used. The Caf1 signal sequence has also been successfully used to export mammalian cytokines (see below). A major problem on recombinant expression in *E. coli* is incorrect folding with accompanying protein degradation or accumulation in an insoluble and inactive form as inclusion bodies.

30

In addition to the *sec* dependent secretion system there are at least two other systems of protein translocation across the bacterial inner membrane. The M13

phage coat protein is also synthesised with an additional SP, but assembly of this protein across the membrane is independent of the *sec* machinery. Recently, a novel pathway involved in secretion of cofactor-containing proteins has been elucidated (Santini, G., et al. (1998) *EMBO Journal*. 17:101-112; Weiner, J., et al. 5 (1998) *Cell*. 93:93-101). Proteins following this pathway have a long leader containing a characteristic 'twin arginine' motif. It is proposed that cofactor attachment occurs in the cytosol and that the fully folded protein is translocated across the inner membrane via products of the *mttABC* operon. In addition, there are cytosolic proteins of *E. coli* which appear to be localised in a privileged site 10 which is sensitive to osmotic shock. Therefore may have some transient access to the periplasm. Such proteins include thioredoxin (Lunn, C.A., and V.P. Pigiet. (1982) *J.Biol.Chem.* 257:11424-11430) involved in disulphide reduction of cellular components, the cytosolic chaperone DnaK (Yaagoubi, A., et al. (1994) *Journal of Bacteriology*. 176:7074-7078), elongation factor Tu (Jacobson, G., et al. 15 (1976) *Biochemistry*. 15:2297-2303) and inner membrane bound components of enterobactin synthase complex (Hantash, F., et al. (1997) *Microbiology*. 143:147-156), and capsule assembly (Rigg, G., et al. (1998) *Microbiology*. 144:2905-2914). It has been suggested that this 'compartment' may be related to transient formation 20 of adhesion zones between the bacterial inner and outer membranes, but nothing is known regarding properties of the protein which targets them to this location nor about the physical nature of this 'compartment'. A number of 'cytosolic' recombinant proteins (i.e. without SP) also behave in a similar manner and are thus presumably targeted to the same cellular location. These include GST fusion proteins, interleukin 1 $\beta$  (Joseph-Liauzun, E., et al. (1990) *Gene*. 86:291-295).

25

## 2. *Extracellular secretion systems*

Six different pathways for export of extracellular proteins have been identified in Gram negative bacteria. Each pathway has been identified in a diverse range of 30 bacteria. The basic properties of these systems are summarised in Fig 1 (recently reviewed by Lory (Lory, S. (1998) *Current Opinion in Microbiology*. 1:27-35).

The Type II pathway, which is considered to be the main terminal branch of the *sec*- dependent pathway, is used for export of many different unrelated soluble proteins. It involves a folded periplasmic intermediate and requires approximately 12 dedicated genes for export across the OM. Alternative terminal branches to the *sec* pathway include specific chaperone-dependent fimbriae assembly and the OM helper pathway. The former pathway also involves a periplasmic intermediate (at least partially folded) but in this case the secreted polypeptide is specifically transported in association with its own chaperone/ outer membrane usher protein system. Outer membrane helpers fold into the outer membrane with concomitant exposure of the effector domain at the cell surface and, in the case of IgA protease, release via self-hydrolysis. Interaction with general periplasmic chaperones e.g. DsbA has been demonstrated as a critical step in secretion pathway for a number of *sec* dependent proteins.

15 Type I, Type III, and most members of Type IV pathways are *sec* independent and mediate secretion of a specific protein (subset of proteins or DNA (Type IV) directly from the cytosol. Type I results in secretion into the external media, whereas Type III targets the secreted protein directly into the eukaryotic cell following contact-stimulated activation of the secretion system. The Type III pathway also shares many features with flagellar assembly systems.

20

### *3. Secretion of recombinant heterologous proteins in Gram-negative bacteria*

Incorrect folding of proteins in the cytosol may lead to degradation or formation of 25 misfolded protein as inclusion bodies. In many instances, therefore, it is desirable to have heterologous expression of recombinant proteins in the bacterial periplasm, at the cell surface, or in the extracellular media, permitting correct folding and formation of a functional product. Proteins secreted to the periplasm of *E. coli* are in an oxidising environment, compared to the reducing environment of the cytosol.

30 The periplasm contains oxidoreductases and chaperones (disulphide bond isomerase, DsbA and C, peptidyl prolyl cis-transisomerase, RotA, SurA, and FkpA) which are essential for the correct folding of proteins (Missiakas, D., and

S.Raina. (1997) *Journal of Bacteriology* 179:2465-2471). In addition, recombinant proteins expressed in the periplasm or secreted to the extracellular medium would represent a high percentage of the final protein content of these respective compartments. Thus, when the final goal is to obtain a purified recombinant product, secretion of the product to the periplasm or externally should greatly facilitate purification protocols. Although there are quite a few systems available for periplasmic localisation of proteins, there is no major system for secretion of extracellular products from *E. coli*. Over the past decade there has also been a great deal of interest in expressing proteins and peptides on the surface of microorganisms. Phage display technology (Winter, G., et al. (1994) *Annual Review of Immunology* 12:433-455) utilises the coat protein of filamentous bacteriophage for surface display of proteins or peptides. Such technology has been applied to the isolation of specific antibody fragments and for the rapid identification of peptide ligands. Interest in surface display in *E. coli* (Georgiou, G., et al. (1993) *Trends in Biotechnology*, 11:6-10) and other Gram negative bacteria has centered around identification of protective epitopes and their applications as live vaccines, production of bacterial adsorbents and whole-cell biocatalysts.

20 Although there has been some success in expressing of proteins, there are a number of limitations within the existing systems as outlined below.

Most secretory/ assembly pathways of *E. coli* have been investigated for their potential exploitation as secretion vehicles for heterologous proteins. These 25 include systems that direct the protein to the periplasm, cell surface or extracellular medium.

### 3.1 SP alone

30 A number of expression vectors use a bacterial SP (often that of the OM protein OmpA) to mediate export across the inner membrane. Destination of the protein depends on the nature of the protein itself. It is not uncommon for proteins

exported in this way in high levels to form insoluble complexes, inclusion bodies, in the periplasm as a result of incomplete folding.

### 3.2. Affinity purification systems

5 Fusion expression systems have been developed to facilitate downstream  
10 purification of recombinant products. Examples include insertion of a His tag for  
15 purification on a Nickel column (Clontech, Qiagen, In vitrogen); fusion to MalE  
20 (New England Biolabs), maltose binding protein, with subsequent purification on  
an amylose column; thioredoxin fusions with PAO (phenyl arsine oxide) resin and  
chitin binding domain fusions with chitin columns (New England Biolabs). By  
inclusion or omission of SP in the vector, some of these systems (e.g. MalE, His  
Tag) can be adapted for periplasmic or cytosolic expression, respectively. In  
general, such vectors contain a highly specific protease cleavage site for  
downstream purification of the product. Fusions functional in both domains, e.g.  
MalE and secreted domain, can be obtained. This, however, is dependent on the  
nature of the protein. The carrier domain may interfere with folding of the  
recombinant protein resulting in protein degradation, insolubility of the protein due  
to membrane association or formation of insoluble inclusion bodies at higher  
concentrations.

### 3.3. Surface display in *E. coli*

25 Insertion of epitopes into major OM proteins (OmpA, LamB, PhoE), flagella, fimbriae. These systems involve insertion of epitopes into a permissive site, i.e. surface loop within OM proteins or flagellar, fimbriae subunits, without affecting assembly of the membrane protein or surface appendage. In general, there are severe size restrictions of the insert (10-60 amino acids) to avoid effects on folding and assembly of the protein. There are reports of surface display of whole 30 proteins by preparing terminal fusions to part of the outer membrane protein, OmpA or of IgA protease. Using a Lpp- OmpA vector, complete enzymes have been localised to the surface of *E. coli* offering the potential of surface display, but

these constructs lead to disruption of the outer membrane with concomitant toxicity to the cell and leakage of periplasmic contents. In addition, the fusion proteins follow the outer membrane protein assembly pathway. This limits the maximum number of surface molecules and more importantly it is evident that 5 completely folded proteins possessing disulphide bonds cannot be assembled across the outer membrane by this route (Klauser, T., et al. (1990) EMBO Journal. 9:1991-1999; Stathopoulos, C., et al. (1996) Applied Microbiology and Biotechnology. 45:112-119).

10 *3.4. Extracellular secretion.*

There have been limited reports on extracellular secretion of unrelated proteins by some of the above mentioned secretion pathways. The Hly Type I secretion pathway has been adapted to delivery of heterologous antigens (Gentschev, I, et al. 15 (1996) Gene 179:133-140). Although apparently successful, this system delivers proteins directly from the cytosol and would preclude any protein which require exposure to the periplasmic space for correct folding, e.g. disulphide bond formation.

20 It is summarised below some of the serious drawbacks associated with recombinant protein expression:

(i) Periplasmic expression systems: Many heterologous polypeptides expressed in *E. coli* are either degraded or form aggregates and inclusion bodies as a result of 25 incorrect folding. This may occur despite targeting of the protein to a preferred location, i.e. the cytosol (with a more reducing environment) or the periplasm (with a more oxidising environment and specific chaperones involved in folding). Employment of a signal sequence to proteins targeted to the periplasm results in varying degrees of efficiency of precursor processing, completion of translocation 30 and correct folding. Some incorrectly folded proteins remain associated with the inner membrane and induce toxicity. In addition, they are extensively degraded resulting in a poor yield. Others accumulate in a non-native conformation as

insoluble aggregates. Systems employing fusion proteins are available. These may to some degree enhance solubility of some recombinant proteins but others remain insoluble due to incomplete folding of the heterologous domain. A system leading to stimulation of the early folding event following translocation across the inner 5 membrane would clearly enable periplasmic expression of many heterologous polypeptides which have thus far eluded successful expression in *E. coli*.

(ii) Surface localisation in Gram negative bacteria: Generally, there is a strict limitation in the size of epitopes which can be expressed at the cell surface using 10 proven surface expression vectors. Systems that permit surface expression of whole domains or proteins by fusion them to an outer membrane protein lead to membrane permeabilisation, periplasmic leakage and toxicity. In addition, there are limitations on the extent to which proteins can be folded if they are to be exported by this pathway. Finally, as these systems all use integral membrane 15 proteins, they are limited with respect to the maximum expression level and would be very laborious to purify.

## SUMMARY OF THE INVENTION

20 The present invention provides bacterial strains for secretion of soluble biologically active recombinant heterologous proteins into periplasm or on a surface/into a cultivation medium of bacteria. The invention exploits the secretion system of Gram-negative bacteria including periplasmic chaperones and usher/secretin proteins. For accomplishing the aim, the bacterial strains 25 simultaneously express the fusion protein (signal peptide)-(mature heterologous protein)-(subunit of a bacterial surface structure, Caf1), periplasmic chaperone specific for the subunit, and outer membrane usher/secretin protein specific for the subunit. Secretion of fusion proteins: (signal peptide of Caf1)-(mature human IL-1 $\beta$ )-(mature Caf1), (signal peptide of Caf1)-(mature human GM-CSF)-(mature Caf1), and (signal peptide of Caf1)-(mature human IL-1ra)-(mature Caf1) that 30 were expressed in *Escherichia coli* simultaneously with the periplasmic chaperone

Caf1M and the usher/secretin protein Caf1A are examples of the use of the invention.

#### BRIEF DESCRIPTION OF THE DRAWINGS

5 Fig. 1. *Extracellular secretion systems in Gram negative bacteria*. Major components include: (i) Type I (e.g. HlyA) - inner membrane (IM) ABC transporter (Hly B and D) + an outer membrane (OM) protein (Tol C), no periplasmic intermediate detected; (ii) Type II (e.g. pullulanase) - Sec system for 10 translocation across IM; translocation across OM - 14 *pul* gene products including an OM secretin (PulC, S), eight IM *pul* products (Pul C, E,F, K-O) and four prepilin-like products (Pul G-J), evidence supports periplasmic intermediate; (iii) 15 Type III (e.g. Yops) - 24 Ysc proteins including an OM secretin (YscC) and IM ATPase (YscN), specific Syc cytosolic chaperones, and Yop B and D required for delivery to eukaryotic cell (iv) Type IV- newly classified group including systems involved in DNA transfer to plant or other bacterial cells (e.g. T-DNA of 20 *Agrobacterium tumefaciens*, 11 *virB* genes, unlikely to go via periplasmic intermediate) and pertussis toxin export (9 *ptl* genes, periplasmic assembly of toxin likely). (v) OM helper (e.g. IgA protease) - Sec system for translocation 25 across IM; no specific accessory proteins; uses a 'helper domain' within the protein which presumably follows the OM protein assembly pathway and folds into the OM exposing the secreted domain on the surface; (vi) Specific chaperone mediated assembly (e.g. Pap pili) - Sec system for translocation across IM; translocation across OM - periplasmic chaperone (PapD) which specifically recognises pilin subunits, OM usher/ secretin protein (PapC) homologous to PulC and YscC of TypeII and TypeIII pathways, respectively.

Fig. 2. Construction of hybrid genes coding for <sup>35</sup>Cafl-hIL-1 $\beta$  fusion proteins.

30 Fig. 3. Expression of <sup>35</sup>Cafl-hIL-1 $\beta$  fusion proteins. A. Coomassie Blue-stained SDS-PAGE of soluble (lanes 2-5) and insoluble (lanes 6-9) proteins from *E. coli* cells transformed with pKKmod (lanes 2, 6), pKKmod/<sup>35</sup>Cafl-hIL-1 $\beta$  (lanes 3, 7),

pKKmod/<sup>s,p</sup>Cafl(-2)hIL-1 $\beta$  (lanes 4, 8), and pKKmod/<sup>s,p</sup>Cafl(+3)hIL-1 $\beta$  (lanes 5, 9). hIL-1 $\beta$  was loaded as a control (lanes 1, 10). B. Immunoblot of the same gel analysed with anti-hIL-1 $\beta$  rabbit polyclonal antibodies. Positions of unprocessed (I) and processed (II) hIL-1 $\beta$  are shown by arrows.

5 Fig. 4. Secretion of <sup>s,p</sup>Cafl-hIL-1 $\beta$  fusion proteins. Immunoblots of periplasmic (A) and soluble cytoplasmic (B) proteins from <sup>s,p</sup>Cafl(+3)hIL-1 $\beta$  (lane 1), <sup>s,p</sup>Cafl(-2)hIL-1 $\beta$  (lane 2), and <sup>s,p</sup>Cafl-hIL-1 $\beta$  (lane 3) expression strains. Corresponding proteins from cells harbouring pKKmod were used for comparison 10 (lane 4). hIL-1 $\beta$  was loaded as a control (lane 5). Proteins were analysed with anti-hIL-1 $\beta$  rabbit polyclonal antibodies. Positions of unprocessed (I) and processed (II) hIL-1 $\beta$  are shown by arrows.

15 Fig. 5. Trypsin digestion of permeabilized cells. Insoluble proteins (lane 1-2) and soluble proteins (lane 3-4) were obtained from <sup>s,p</sup>Cafl(-2)hIL-1 $\beta$  expression cells before (lane 1,3) and after (lane 2,4) trypsin treatment.

20 Fig. 6. Construction of pCIC plasmid coding for the <sup>s,p</sup>Cafl(-2)-hIL-1 $\beta$ -Cafl fusion protein.

Fig. 7. The N-terminal sequence of mature CIC. After partial purification of 25 periplasmic fractions on a DEAE-Sepharose CL-6B column (Pharmacia, Sweden) proteins were separated by SDS-PAGE followed by blotting onto a PVDF-membrane (Amersham, UK). The desired bands were excised and placed onto a polybrene coated and precycled glass fiber filter. Amino acid sequence analyses were performed with an Applied Biosystems model 477A protein sequencer equipped with on-line Applied Biosystems model 120A phenylthiohydantoin amino acid analyser.

30 Fig. 8. CaflM facilitates expression of <sup>s,p</sup>Cafl(-2)-hIL-1 $\beta$ -Cafl (CIC). A. Coomassie Blue-stained SDS-PAGE of periplasmic fraction from cells

transformed with: pCIC (lane 1), pFMA (lane 2), pMA-CIC (lane 3), pA-CIC (lane 4), pM-CIC (lane 5), pMA-PrCIC (lane 6), pA-PrCIC (lane 7), pM-PrCIC (lane 8), pCIC, pCaf1M (lane 9), pCIC, and pCaf1MA (lane 10). B. Immunoblot of the same gel analysed with anti-hIL-1 $\beta$  rabbit polyclonal antibodies.

5

Fig. 9. Graphical representation of plasmids constructed for operon-like co-expression experiments with Caf1M, Caf1A, and CIC. pFMA and pCIC are given for comparison.

10 Fig. 10. Detection of CIC in periplasm with monoclonal anti-IL antibodies using ELISA. Detection of CIC in dilutions of a periplasm aliquot from control cells carrying pTrc99 with plasmid (triangles), from cells carrying pCIC (circles) and from cells carrying both pCIC and pCaf1MA (squares).

15 Fig. 11. Construction of pFGMF1 and pFRF275 plasmids coding for the scaf-GMCSF-Caf1 and scaf-IL1ra-Caf1 fusion protein.

Fig. 12. Fractionation of proteins expressed in JM101 cells with plasmids pFGMF1 (lanes 3,4,7,8), pFRF275 (lanes 1,2,5,6), and pCaf1M (lanes 2,4,6,8).  
20 After induction cells were precipitated, resuspended in buffer with lysozyme and incubated on ice for 1 hour followed by sonication for 1 min. Soluble and insoluble proteins were separated by centrifugation. Lanes 1-4 - insoluble fraction. lanes 5-8 - soluble fraction.

25 Fig. 13. Expression of scaf1-GMCSF-Caf1 fusion gene. A. SDS-PAGE analysis of periplasmic proteins obtained from cells harboring pFGMF1 (lane 1), pFGMF1 and pCaf1M (lane 2), and pFGM13 (lane 3). B. Protein immunoblot analyzed with anti-GMCSF rabbit polyclonal antibodies. Lanes 1,2 - total cell proteins, lanes 3-5 - periplasmic proteins. Protein fractions were obtained from cells harboring pFGMF1 (lane 1,4), pFGMF1 and pCaf1M (lane 2,3), and pFGM13 (lane 5). C. Immunoblot of periplasmic proteins analyzed with anti-Caf1 rabbit polyclonal

antibodies. Cells produced GMCSF-Caf1 fusion protein without (lane 1) or with (lane 2) the Caf1 M chaperon.

Fig. 14. Expression of scaf1-IL1ra-Caf1 fusion gene. A. SDS-PAGE of 5 periplasmic proteins obtained from cells harbouring pFRF275 alone (lane 1) or with pCaf1M (lane 2). B. Immunoblot of the same gel analysed by anti-IL1ra goat polyclonal antibodies.

#### DETAILED DESCRIPTION OF THE INVENTION

10

According to the present invention, secretion of proteins by gram-negative bacteria can be accomplished with adhesins via chaperone/usher pathway. The periplasmic chaperones of the system:

- 15 • mediate the partitioning of nascently translocated subunits out of the inner membrane and into the periplasm;
- help in the folding of nascent subunits into a native conformation;
- protect the subunits from a proteolytic degradation.

20 The periplasm of wild type strains of bacteria contains oxidoreductases and additional chaperones (disulphide bond isomerase, DsbA and C, peptidyl prolyl cis-transisomerase, RotA, SurA and FkpA) which are essential to the correct folding of proteins.

25 The *caf* operon is the simplest operon in comparison to the size of the operons dedicated to most of the other extracellular secretion systems (Karlyshev, A.V. et al. (1994) in Biological Membranes: Structure, Biogenesis and Dynamic. NATO-ASI Series, vol. H-82, Op den Kamp, J.A.F., ed., pp. 321-330, Springer-Verlag. Berlin). It comprises of *caf1R* encoding a transcription regulator, the structural 30 gene for the Caf1 polypeptide (Galyov, E.E. et al. (1990) FEBS Lett. 277, 230-232) and two genes encoding products for specific secretion of Caf1 - the Caf1M and periplasmic chaperone (Galyov, E.E. et al. (1991) FEBS Lett. 286, 79-82) and

Caf1A outer membrane protein (Karlyshev, A.V. et al. (1992) FEBS Lett. 297, 77-80).

5 Caf1 polypeptide is synthesised with a 20 aa cleavable SP. Following translocation across the inner membrane (presumably via the *sec* pathway), Caf1M binds to mature Caf1 and protects it from proteolytic degradation by assisting its folding and also its release from the inner membrane (Zav'yalov, V. et al. (1997) Biochem. J., 324, 571-578; Chapman D. et al. (1999) J. Bacteriology, in press).  
10 Interaction of subunit with Caf1M chaperone then stimulates a signal for interaction with the OM protein Caf1A and Caf1A translocates Caf1 to the cell surface. Surface assembled Caf1 forms an amorphous capsule-like structure on the surface of the bacteria. This surface structure can be readily recovered with the bacterial cells and washed off the surface to give a relatively pure preparation of the Caf1 protein. Caf1M chaperone and Caf1A usher are specific for Caf1 subunit  
15 and not suitable for a secretion of heterologous proteins from this point of view. However, surprisingly it was found in the present invention that the fusions of three different heterologous proteins with the Caf1 subunit are secreted in soluble biologically active form and protected from proteolytic degradation when expressed simultaneously with the Caf1M chaperone.

20

According to the present invention Caf system can be applied for:

- production of recombinant heterologous proteins in the periplasm of bacteria in a soluble, biologically active conformation;
- 25 • surface display of whole recombinant heterologous proteins on Gram-negative bacteria - for obtaining heterologous live attenuated vaccines, construction whole cell ligands, whole cell biocatalysts, purification of recombinant heterologous protein from cell surface by washing fusion protein out of the cell surface, or by proteolytic cleavage of fusion protein;
- 30 • surface display of heterologous amino acid sequences (epitopes) within Caf1 subunit, and use as ligand, vaccine candidate etc.

*Potential applications of the caf system:*

1. *CafISP-export across the IM:* As with many other secretory proteins the Caf1 signal sequence can be used directly for secretion of proteins across the bacterial inner membrane. It contains no positively charged residues at its C-terminus that might interfere with export and can therefore be used directly. As the secreted protein would follow the Sec pathway to the periplasm, the efficiency of export of the heterologous protein would also depend on the nature of the protein to be secreted. In particular, efficiency of secretion would depend on the N-terminus of the mature polypeptide, absence of long potential anchor sequences within the secreted protein, and absence of protein folding prior to export. The first two features can be addressed by altering the relevant sequences. 'Premature' folding is more difficult to control although excess of cytosolic SecB, GroEL, DnaK chaperone may delay folding in some instances. Following translocation across the IM and SP cleavage, release of the heterologous protein from the IM and folding of the protein would again be dependent on the nature of the heterologous protein itself.
2. *CafIM chaperoning in the periplasm:* The *caf* system has the added advantage that it includes a periplasmic chaperone which specifically binds to the Caf1 subunit. Caf1M prevents degradation of the Caf1 subunit, probably via enhanced folding and release from the IM. Chaperone recognition is via the C-terminus of the Caf1 subunit although interaction with other parts of the Caf1 subunit may also be required for high affinity binding. Hence, fusion of the Caf1 subunit C-terminally to the protein destined for release may stimulate folding of the heterologous protein, aid release of the protein from the IM, increase its solubility and prevent aggregation (formation of periplasmic inclusion bodies) and proteolytic degradation.
3. *Caf1A - export of heterologous proteins extracellularly across the outer membrane.* Caf1M chaperone also targets Caf1 to the OM secretin Caf1A. Following interaction of the complex with Caf1A secretin, the Caf1 is translocated across the outer membrane and forms a large polymeric structure

on the cell surface -anchored via CaflA. For the case that the anchoring is disfavored a proper mutant of Cafl can be prepared. Thus, inclusion of CaflA in the expression vector would permit targeting of the recombinant protein to the CaflA secretin.

5 4. Hybrid proteins exported by this system could be proteolytically cleaved to remove Cafl by introduction of an appropriate cleavage site between Cafl and the protein of interest. For proteins arrayed on the cell surface of *E. coli* this provides an enormously efficient method of purification from isolated cells. Cell surface localised hybrid could also be used for vaccine preparation, ligand binding or as whole cell biocatalyst. Proteins secreted from the cell or 10 recovered from the periplasm would be already in a highly purified form. All proteins secreted by this system would be exposed to the periplasm, offering the milieu for correct folding of secreted proteins. Using this system, the probability of folding into a functional protein/ domain would be further 15 increased for some fusion proteins due to interaction with the CaflM chaperone.

In addition to secretion of whole proteins or domains as an N-terminal fusion to the Cafl subunit (or possibly smaller domain thereof), it will undoubtedly also 20 prove to be the case that short epitopes can be inserted into permissive sites within the Cafl subunit permitting surface display of epitopes. Similarly, short peptides can be inserted into the FGL site of the CaflM chaperone.

25 *Summary of potential advantages of the caf system as applied to heterologous secretion in Gram negative bacteria:*

- despite the fact that protein hIL-1 $\beta$  is insoluble in the periplasm and a fusion hIL-1 $\beta$ -Cafl is also insoluble, co-expression with the CaflM chaperone surprisingly led to a soluble product. Hence, the CaflM chaperone catalyses 30 folding, not only of the carrier Cafl domain, but also of the IL-1 $\beta$  domain, resulting in release from the membrane and formation of a processed, soluble product. Because of correct folding the heterologous fusion protein is not

degraded and accumulates in high levels. *I.e. Caf1M can stimulate folding, membrane release and solubility of heterologous polypeptides expressed as Caf1 fusion proteins;*

- while available systems of surface display of whole proteins in *E. coli* result in problems of loss of integrity of the outer membrane and have limitations as to the extent of folding of the exported protein, the caf system has not these limitations. The fusion protein is anchored to the cell surface or released as a soluble protein without disrupting the membrane. Similarly, assembly does not follow the folding pathway of outer membrane proteins and a fully folded fusion protein is exported;
- option to readily express high levels of recombinant protein in periplasm (Caf1M alone) or cell surface/ media (Caf1M + Caf1) using the same system;
- ready purification from the cell surface by proteolytic cleavage or recovery in high levels from periplasmic fractions (whichever preferred);
- exquisitely simple compared to other potential *extracellular* secretion vectors employing only a periplasmic chaperone (Caf1M) and outer membrane protein (Caf1A);
- ability to secrete extracellularly a whole protein or part thereof
- transient or permanent periplasmic localisation of fusion protein permits correct folding and oxidation of secretory proteins.

The invention is further illustrated here below with specific non-limiting examples. The examples describe the secretion of the fusion proteins (signal peptide of Caf1)-(mature human IL-1 $\beta$ )-(mature Caf1), (signal peptide of Caf1)-(mature GM-CSF)-(mature human Caf1) and (signal peptide of Caf1)-(mature human IL-1 $\alpha$ )-(mature Caf1) expressed in *Escherichia coli* simultaneously with the periplasmic chaperone Caf1M and the usher/secretin protein Caf1A as the examples of the use of the system. Even though Examples describe secretion of 3 proteins it is obvious that any other protein can be secreted using this system. In addition, although *E.coli* is used as the host microbe containing the expression system in these

examples, it is obvious that also other cells having proper periplasmic space for accommodating the caf expression system can be used.

#### EXAMPLE 1

##### 5 **Fusion of hIL-1 $\beta$ with the signal peptide of the Caf1 protein.**

The Caf1 signal peptide (<sup>s,p</sup>Caf1) was fused with hIL-1 $\beta$  in three variants (Fig.2).

1. <sup>s,p</sup>Caf1-hIL-1 $\beta$  was the straight fusion where the last amino acid of the Caf1 signal peptide was joint to the first amino acid of hIL-1 $\beta$ .
2. <sup>s,p</sup>Caf1(-2)-hIL-1 $\beta$  was the straight fusion with a mutation in the Caf1 signal peptide Asn(-2)Asp.
- 10 3. <sup>s,p</sup>Caf1(+3)hIL-1 $\beta$  was the fusion containing in the joint region three N-terminal amino acids from mature Caf1 to preserve the natural processing site.

15 Here and thereafter DNA manipulations and transformation of *E. coli* were accomplished according to Maniatis, T., et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor. Restriction enzymes, mung-bean nuclease, and T4 DNA ligase were purchased from Promega (USA). Taq DNA polymerase (Hytest, Finland) was used for polymerase chain reaction (PCR) experiments. Nucleotide sequencing 20 was carried out using the TaqTrack sequencing kit (Promega, USA). Elution of DNA fragments from agarose gels was performed with the USBClean kit (USB, USA). Following oligonucleotides were used in EXAMPLES 1 and 2:

CAF-RI	5'-GGGAATTCAAGAGGTAATATATGAAAAAAATC-3'
IL-PST	5'-CCGCCTGCAGATCGGGCACCTGTACGATCACTG-3'
CAF-PST	5'-CCGCCTGCAGTTGCAATAGTTCAAATA-3'
IL-Primer	5'-AGAACACCACTTGTGCTCC-3'
Blunt	5'-TGGAACTATTGCAACTGCAAATCGGGCACCTGTACGA-3'
3AA	5'-GCAACTGCAAATCGGGCAGATTAGCACCTGTACGATC- ACTG-3'
IL-BamHI	5'-ACCGGATCCACCTCCACCAGATCCACCTCCGGAAGACACA

	AATTGCATGG-3'
BamHI-	5'-GGTGGATCCGGTGGTGGATCTGCAGATTTAACTGCA
Caf	AGCAC-3'
Caf-Sall	5'-GCCAAGCTTGTGACGAGGGTTAGGCTCAAAGT-3'
SBEKP-1	5'-TCGACAGATCTGAATTCCGGTACCGGCTGCA-3'
SBEKP-2	3'-GTCTAGAGCTTAAGGCCATGGCCG-5'
STOP	5'-GATCATTAAATTAAT-3'
TRC	5'-CCAGATCTGGCAAATATTCTGAAATG-3'

The genetic constructs coding three fusion proteins <sup>s.p.</sup>Caf1-hIL-1 $\beta$  were made according to the scheme in Fig. 2. The *Eco*RI-*Pst*I fragment (about 110 bp) coding 5 the Caf1 5'-untranslated region and N-terminal part of Caf1 signal peptide with the mutation Asn(-2)Asp was obtained by PCR using CAF-RI and IL-Pst oligonucleotides and the template plasmid pKM4 (Karlyshev, A.V., et al. (1992) FEBS Letters 297, 77-80) followed by digestion of PCR1 product with *Eco*RI and *Pst*I. The *Pst*I-*Hind*III fragment coding for the C-terminal part of Caf1 signal 10 peptide joint to the N-terminus of hIL-1 $\beta$  was obtained by PCR where CAF-PST and IL-Primer oligonucleotides and the template plasmid pPR-TGATG-hIL-1 $\beta$ -tsr (Mashko, S.V., et al. (1991) Gene 97, 259-266) were used. The PCR2 product was digested with *Pst*I and *Hind*III. The two fragments were ligated together with a 15 vector fragment pUC19/*Eco*RI-*Hind*III. The nucleotide structure of the resulted *Eco*RI-*Hind*III insert was verified by DNA sequencing. The *Eco*RI-*Hind*III fragment obtained as described and the *Hind*III-*Bam*HI fragment isolated from pPR-TGATG-hIL-1 $\beta$ -tsr were ligated together with the *Eco*RI-*Bam*HI vector 20 fragment of pUC19 $\Delta$ *Hind*III (a derivative of pUC19 where the *Hind*III site was deleted by filling of sticky ends followed by blunt-end ligation). The *Eco*RI-*Bam*HI fragment thus obtained encoded the <sup>s.p.</sup>Caf1(-2)hIL-1 $\beta$  fusion protein. The point mutation G to A converting the <sup>s.p.</sup>Caf1(-2)hIL-1 $\beta$  gene into the Caf1-hIL-1 $\beta$  gene was made by means of two-step PCR procedure (Landt, O., Grunert, H.-P. and Hahn, U. (1990) Gene 96, 125-128) using the BLUNT oligonucleotide and two flanking primers (M13 Sequence Primer and IL-Primer) with

pUC19Δ*Hind*III/<sup>s,p</sup>*Caf*l(-2)hIL-1 $\beta$  plasmid as a template. In the first step, an intermediate PCR product was obtained from the mutagenic BLUNT oligonucleotide and the M13 Sequence Primer. After purification from an agarose gel, the intermediate PCR product was used with the IL-Primer in the second PCR 5 step. The latter PCR product was digested with *Eco*RI and *Hind*III followed by ligation into the vector fragment made by restriction of the pUC19Δ*Hind*III/<sup>s,p</sup>*Caf*l(-2)hIL-1 $\beta$  plasmid with the same enzymes. Thus, the *Eco*RI-*Hind*III fragment providing the Asn(-2)Asp mutation in the <sup>s,p</sup>*Caf*l(-2)hIL-1 $\beta$  fusion protein was replaced by the corresponding fragment coding the natural 10 *Caf*l signal peptide, and the <sup>s,p</sup>*Caf*l-hIL-1 $\beta$  gene was obtained. The <sup>s,p</sup>*Caf*l(+3)hIL-1 $\beta$  gene was constructed in a similar way using the 3AA oligonucleotide as a mutagenic primer and pUC19Δ*Hind*III/<sup>s,p</sup>*Caf*l-hIL-1 $\beta$  plasmid as a template. In all three hybrid genes the *Eco*RI-*Hind*III fragments were sequenced to prove correct nucleotide structures.

15 *Eco*RI-*Bam*HI fragments coding for the fusion proteins were excised from plasmids described above and were transferred into pKKmod (a derivative of pKK 223-3 (Pharmacia) created by deletion of *Bam*HI- and *Sall*-sites in the *tet* gene region and by replacement of the pKK 223-3 polylinker with the pUC18 20 polylinker). The expression plasmids pKKmod/<sup>s,p</sup>*Caf*l(-2)hIL-1 $\beta$ , pKKmod/<sup>s,p</sup>*Caf*l-hIL-1 $\beta$  and pKKmod/<sup>s,p</sup>*Caf*l(+3)hIL-1 $\beta$  were thus obtained, and *E. coli* JM105 [F' *tra*D36, *lac*<sup>F</sup>Δ(*lac*Z)M15, *pro*A<sup>+</sup>*B*<sup>+</sup>/*thi*, *rpsL*(Str<sup>+</sup>), *endA*, *sbcB*, *sbcC*, *hsdR4*(r<sub>k</sub><sup>-</sup> m<sub>k</sub><sup>+</sup>), Δ(*lac*-*pro**AB*)] (NE BioLabs) were transformed with these plasmids.

25 Expression and processing of <sup>s,p</sup>*Caf*l-hIL-1 $\beta$ , <sup>s,p</sup>*Caf*l(-2)hIL-1 $\beta$ , and <sup>s,p</sup>*Caf*l(+3)hIL-1 $\beta$  fusion proteins were monitored by SDS-PAGE electrophoresis and immunoblotting. The recombinant *E. coli* strains were grown to about 0.5 AU at 600 nm, IPTG (0.5 mM) was added and cells were grown further. Soluble and 30 insoluble proteins were analysed as separate fractions obtained as follows. After certain time intervals, samples of the growth culture with fixed amount of cells

were withdrawn. cells were precipitated and washed in 25 mM Tris·Cl, pH 7.5/100 mM NaCl/1 mM EDTA. Reprecipitated cells were suspended in 50 mM H<sub>3</sub>PO<sub>4</sub>-Tris, pH 6.8 and sonicated using Labsonic U Generator (B. Braun Diessel Biotech). The sonicated samples were centrifuged at 14,000 g for 10 min. 5 Supernatant contained soluble proteins. Insoluble proteins were extracted from the pellet with the SDS-PAGE sample buffer containing 2% SDS and 5% β-mercaptoethanol.

Here and thereafter 10-15% SDS-PAGE of the proteins from both soluble and 10 insoluble fractions was performed according to the Laemmli procedure in Mini-PROTEAN II apparatus (BIO-RAD). Proteins were transferred to a Hybond-C membrane (Amersham) by electroblotting in mini-Trans-Blot Electrophoretic Transfer Cell (BIO-RAD) followed by immunodetection. Results of immunoblots were revealed with the ECL kit (Amersham). Polyclonal rabbit antibodies to hIL- 15 1β (Calbiochem) and monospecific antibodies to Cafl were used for visualisation of IL-1β- and Cafl-containing fusion proteins. Rabbit polyclonal antibodies to CaflM and mouse polyclonal antibodies to CaflA were used for detection CaflM and CaflA, correspondingly. The binding of the primary antibodies were visualised by rabbit (Calbiochem) and mouse (Amersham) conjugate.

20 High expression was observed for all three fusion proteins (about 20-30% of total cell protein). However, in all three recombinant strains fractions of soluble proteins did not contain significant amounts of recombinant proteins (Fig. 3A and B, lanes 1-4). The recombinant hIL-1β protein was discovered in fractions of 25 insoluble cell proteins (Fig. 3A and B, lanes 6-9). An additional hIL-1β containing protein was clearly seen in all three constructs. Most probably, it was a product of rather specific proteolysis of unprocessed fusion protein accumulated in cytoplasm.

30 The straight fusion appeared not to be processed (Fig. 3A and B, lanes 7), while in the other two products the removal of Cafl signal peptides from hIL-1β moieties achieved considerable levels (Fig. 3A and B, lanes 8,9).

To elucidate secretion of processed recombinant hIL-1 $\beta$  proteins, periplasmic proteins were separated from cytoplasmic proteins by osmotic shock procedure as follows. Cells precipitated from 10 ml of growth medium were suspended in 200 5  $\mu$ l 20% (w/v) sucrose/ 0.3 M Tris-HCl, pH 8.0/ 0.5 mM EDTA and kept at room temperature for 10 min. Sucrose-treated cells obtained by centrifugation were suspended in ice-cold 10 mM MgCl<sub>2</sub>/ 0.1 mM PMSF and incubated in ice water for 10 min. After centrifugation, periplasmic proteins were recovered in supernatant fractions. The pellets were suspended in 50 mM H<sub>3</sub>PO<sub>4</sub>-Tris, pH 6.8 10 and sonicated followed by centrifugation at 14,000 g for 10 min. These supernatants contained soluble cytoplasmic proteins. The activity of the cytoplasmic enzyme, glucose 6-phosphate dehydrogenase, was checked as a control of the purity of the periplasmic fraction (Naglak T. J. et al. 1990, 12, 603- 15 611). The obtained samples of the periplasm fraction did not show more than 0.25 % of the cytoplasmic glucose 6-phosphate dehydrogenase activity.

Western blot analysis of periplasmic and soluble cytoplasmic proteins (Fig. 4) has shown that only minor parts of the processed fusion proteins were found in a soluble form in periplasms of <sup>s.p.</sup>Caf1(-2)hIL-1 $\beta$  and <sup>s.p.</sup>Caf1(+3)hIL-1 $\beta$  expression 20 strains (Fig. 4A, lanes 1,2). No processed protein was detected in periplasm of <sup>s.p.</sup>Caf1-hIL-1 $\beta$  expression strain (Fig. 4A, lane 3). Processed hIL-1 $\beta$  proteins were absent in cytoplasmic fractions (Fig. 4B).

The data obtained demonstrated that the <sup>s.p.</sup>Caf1(-2)hIL-1 $\beta$  and <sup>s.p.</sup>Caf1(+3)hIL-1 $\beta$  25 fusion proteins were partly processed (in contrast to the Caf1-hIL-1 $\beta$  fusion protein). Processed products were secreted into periplasm. However, the majority of the processed recombinant hIL-1 $\beta$  accumulated in an insoluble form. Moreover, this insoluble protein is hidden from liquid phase of the periplasm since it is not digested during trypsin treatment of permeabilized cells (Fig. 5). Variation of 30 growth temperature and concentration of the inductor did not facilitate the Caf1 signal peptide removal and secretion of hIL-1 $\beta$  (data not shown).

## EXAMPLE 2

**Fusion of <sup>s.p.</sup>Caf1(-2)hIL-1 $\beta$  with the Caf1 protein sequence and expression of fused protein in the presence of chaperone and usher proteins.**

5

The CIC (<sup>s.p.</sup>Caf-IL-1 $\beta$ -Caf) protein was designed where the <sup>s.p.</sup>Caf1(-2)hIL-1 $\beta$  amino acid sequence (N-terminus of the fusion) was linked to the Caf1 protein sequence (C-terminus of the fusion) through a spacer GlyGlyGlySer repeated three times. The spacer was inserted to minimise possible conformational 10 problems of the two proteins fused in CIC.

The pCIC expression vector was constructed according to the scheme on Fig. 6. The IL part of CIC was obtained by PCR using IL-Pst and IL-BamHI oligonucleotides as primers and the pUC19 $\Delta$ HindIII/ <sup>s.p.</sup>Caf1(-2)hIL-1 $\beta$  plasmid as 15 a template. The Caf1 part of CIC was obtained by PCR using BamHI-Caf1 and Caf1-SalI oligonucleotides as primers and the pKM4 plasmid as a template. PCR products were digested with restriction enzymes as shown in Fig. 3 followed by triple ligation with the pUC19 $\Delta$ HindIII/ <sup>s.p.</sup>Caf1(-2)hIL-1 $\beta$  vector obtained by digestion with *Pst*I and *Sal*I. To produce pCIC, *Eco*RI-*Sal*I fragment was excised from 20 pUC19 $\Delta$ HindIII/CIC plasmid and ligated into the *Eco*RI-*Sal*I vector obtained from pTrc99 $\Delta$ Nco plasmid (a derivative of pTrc99a (Pharmacia) created by digestion with *Nco*I followed by mung-bean nuclease treatment and ligation of blunt ends).

After IPTG induction, *E. coli* JM105 cells harbouring pCIC produced the CIC 25 fusion protein that was successfully processed. A precise removal of the signal peptide was proved by N-terminal sequencing of the soluble periplasmic CIC (Fig. 7). However, only a minor part of the mature IL-1 $\beta$ -Caf1 protein was extracted by an osmotic shock procedure (Fig. 8B, lane 1). Main part of the protein was found in the membrane fraction (data not shown), similar to mature hIL-1 $\beta$ .

Expression/secretion of CIC in the presence of the chaperone (Caf1M) and the usher (Caf1A) was studied in two systems. a) *caf1m* and *Caf1a* genes were situated on a pACYC plasmid compatible with the pTrc99 derivative coding for CIC. b) *caf1m*, *caf1a* and *cic* genes together formed an artificial operon.

5

In compatible plasmid experiments, *E. coli* JM105 cells were transformed simultaneously with pCIC and pCaf1M (a plasmid carrying *caf1m* gene under *tac* promoter in a pACYC184 derivative, see EXAMPLE 3) to study influence of the chaperone on secretion of hIL-1 $\beta$  fusion protein. Further, pCaf1MA, a Caf1M-chaperone expression/secretion plasmid, was created as follows. *Apa*LI-*Apa*LI fragment containing *caf1m* and *caf1a* genes under *trc* promoter was excised from the pFMA plasmid (Chapman D., et al. Structural and functional significance of the FGL sequence of the periplasmic chaperone, Caf1M, of *Yersinia pestis* J. Bacteriology, *in press*) and ligated into the vector obtained by digestion of pCaf1M with *Apa*LI. *E. coli* JM105 cells were transformed simultaneously with pCIC and pCaf1MA to obtain co-expression of all three proteins.

20

Several plasmids were constructed to study expression and secretion of CIC, Caf1M and Caf1A proteins produced from an artificial operon (Fig. 9). To replace the *caf1* gene with a SBEKP synthetic polylinker, a pMA-link plasmid was obtained by triple ligation of a pFMA/*Pst*I-*Spe*I vector, a *Spe*I-*Pst*I fragment of pFMA, and SBEKP-1 and SBEKP-2 oligonucleotides annealed together. A *Sal*I-*Sal*I fragment coding for the Caf1A protein followed by self-ligation of the vector. A pA-link plasmid was obtained from the pMA-link plasmid by excision of a *Bam*HI-*Bam*HI fragment coding for the C-terminal part of Caf1M protein. To interrupt the Caf1M translation frame, the stop-codon was inserted by ligation of a self-complementary STOP oligonucleotide into the *Bam*HI site. The insertion of the STOP oligonucleotide resulted in the loss of *Bam*HI site. A fragment coding for the CIC protein was excised from pCIC with *Eco*RI and *Sal*I, cloned in pBCSK<sup>+</sup> (Stratagene, USA), and recovered with *Eco*RI and *Kpn*I. To obtain pMA-CIC, pM-CIC, and pA-CIC plasmids, the *Eco*RI-*Kpn*I fragment was cloned into

corresponding sites of pMA-link, pM-link, and pA-link. A plasmid pM-PrCIC differed from pM-CIC by the presence of additional *trc* promoter before the CIC gene, and was obtained as follows. A DNA fragment coding for the *trc*-promoter and the 5'-region of the *cic* gene was obtained by PCR using TRC and CAF-Pst 5 oligonucleotides as primers and pCIC as a template. The PCR product was digested with *Bgl*II and *Eco*RI followed by ligation of the *Bgl*II-*Eco*RI fragment coding for the *trc*-promoter into corresponding sites of pM-CIC. pMA-PrCIC and pA-PrCIC were obtained by ligation of the *Bgl*II-*Kpn*I fragment from pM-PrCIC 10 into corresponding sites of pMA-link and pA-link. *E. coli* NM522 [F', *proAB*, *lacI*<sup>R</sup>*Δ(lacZ)M15/ supE*, *thi-1*, *Δ(lac-proAB)*, *Δ(hsdSM-mcrB)5*, (*r<sub>k</sub>* *m<sub>k</sub>*)] (Stratagene, USA) was used as a host strain for plasmids described here.

Similar results were obtained in both co-expression methods. Simultaneous 15 expression of CIC and Caf1M significantly increased the concentration of the mature CIC in the periplasm. The results demonstrated that periplasmic molecular chaperone promoted a release of the secreted protein from the inner membrane as well as prevented protein from degradation and unspecific aggregation.

In the presence of Caf1M a significant part of the mature CIC protein was detected 20 in a periplasmic fraction (Fig. 8B, lanes 3,4,6,7,9,10). Elevated amount of secreted and soluble CIC co-expressed with Caf1M demonstrates that the periplasmic molecular chaperone promotes a release of the secreted protein from an inner membrane. The critical role of the Caf1 part in the CIC protein in this promotion was proved in a similar experiment with a plasmid differed from pCIC by deletion 25 of one base pair in the spacer. The deletion resulted in the frame shift of the Caf part of the fused gene. No facilitation of hIL-1 $\beta$  secretion was observed when cells harboured the mutated plasmid (data not shown).

Caf1M prevented CIC from degradation. As shown in Fig. 8B, lane 1, the mature 30 CIC protein in the periplasmic fraction rapidly degraded into a truncated form with a molecular weight about 20-23 kDa. In the presence of Caf1M the truncated form disappeared (Fig. 8B, lanes 3,4,6,7,10). The amount of the degraded form

correlated with the amount of Caf1M. For example, when Caf1M was expressed from pCaf1M at a low level the truncated form was found (Fig. 8B, lane 9). However, the Caf1M amount was sufficient to facilitate a release of the fusion protein from the membrane. The specific digestion of the mature CIC occurred at a 5 site in the Caf1 part of the fusion protein since the truncated form was well detected by the IL antibodies but not detected by the Caf1 antibodies (data not shown). Most probably the cleavage was due to the action of protease DegP, which had been shown is induced by unfolded secreted capsular or pilus subunits and cleaves them (Soto, G.E., et al. (1998) EMBO J., 17, 6155-6167).

10

The hIL-1 $\beta$  part of CIC secreted in the presence of Caf1M was correctly folded. The mature CIC protein was detected with monoclonal antibodies to hIL-1 $\beta$  in ELISA (Fig. 10). ELISA was performed using monoclonal mouse antibodies to hIL-1 $\beta$  (HyTest, Finland) as it was described previously (Zav'yalov, V., et al. 15 (1997) Biochem. J., 324, 571-578).

CIC was excreted onto a cell surface when expressed with both Caf1M and Caf1A. The presence of the molecular chaperone and the usher protein in cells transformed with pMA-CIC or with pCIC and pCaf1MA together was proved by 20 immunoblotting with anti-Caf1M and anti-Caf1A antibodies (data not shown). To prove the excretion of CIC onto a cell surface, cell agglutination experiments were performed with reticulocyte monoclonal diagnosticum for detection of *Yersinia pestis* (Middle Asian Reaserch Institute, USSR). The *E. coli* cells expressed CIC, Caf1M, Caf1A were able to precipitate reticulocytes with surface bound 25 monoclonal antibody to Caf1 at the concentration about 10<sup>7</sup> cells/ml. Moreover, according to ELISA performed with anti-hIL-1 $\beta$  monoclonal antibodies, a small amount of hIL-1 $\beta$  was detected in cultural medium.

### EXAMPLE 3

30 **Co-expression of hGMCSF-Caf1 fusion protein and the chaperone**

The expression plasmid for the hGM-CSF-Cafl fusion protein was based on pFGM13 described in: Petrovskaya L.E. et al. (1995) Russian Journal of Bio-organic Chemistry, 21, 785-791. The pFGM13 plasmid contained a synthetic gene coding for hGMCSF with Cafl signal sequence (scaf1) cloned into *Hind*III and 5 *Xba*I sites of pUC19. Transcription was controlled by the *lac* promoter and was inducible with IPTG. Translation of the *scaf1-gmcsf* gene in pFGM13 initiated at the first methionine codon of *lacZ* gene and utilises its Shine-Dalgarno sequence. A primary structure of the signal sequence (scaf1) encoded by this gene differed from wild-type one at its N-terminus, which contained seven extra residues from 10  $\beta$ -galactosidase N-terminus.

The construction of the pFGMF1 plasmid coding for the *gmcsf-cafl* gene is shown in Fig. 11. The *scaf1-gmcsf* gene of pFGM13 was modified for the following cloning of the fragment from pCIC containing a spacer (4GlySer)<sub>3</sub> and the Cafl coding region. The *Kpn*2I site was introduced at 3'-terminus of the *gmcsf* gene by 15 PCR using two primers (5'ATCGGAAATGTTGACCTTCAAG and 5'ATTATTCCGGACTCCTGCACTGGTTCCCAGC) and pFGM13 as a template. Pfu DNA Polymerase was used for a maximal fidelity. The PCR fragment was treated with *Kpn*2I followed by ligation into the *Eco*RV-*Sal*GI large 20 fragment of pFGM13 together with *Kpn*2I-*Sal*GI fragment from pCIC. The final plasmid (pFGMF1) contained the gene encoding a hybrid precursor consisting of the scaf1 signal sequence, GMCSF with 2 N-terminal amino acid changes (Ala<sub>2</sub>Pro<sub>3</sub> to Asp), a Ser(4GlySer)<sub>3</sub> spacer, and Cafl. The plasmid structure was confirmed by restriction analysis and sequencing of amplified regions.

25 Expression of the hybrid precursor gene in JM101 *E. coli* cells harbouring pFGMF1 was induced with 0.2 mM IPTG after cell culture reached an optical density 0,5-0,8. The growth with IPTG continued for 3 h. Cells from 1 ml were collected by centrifugation and the pellet (total cell protein sample) was analysed 30 by SDS-PAGE electrophoresis. A fraction of periplasmic proteins was isolated by the cold osmotic shock procedure (see EXAMPLE 1) from the cell pellet obtained from the rest of the culture.

SDS-PAGE analysis of the total cell protein sample from *E. coli* cells harbouring pFGMF1 revealed the presence of a large amount of the fusion protein with molecular mass corresponding to the hGMCSF-Caf1 fusion protein (31 kDa).

5 When the cells were destroyed by sonication this protein was localised in an insoluble fraction (Fig. 12, lane 3, 4).

Examination of periplasmic extracts by SDS-PAGE and Western blotting has shown that some hGMCSF-Caf1 fusion was translocated to the periplasm. Its 10 electrophoretic mobility was the same as of the insoluble protein. The fusion interacts with both anti-hGMCSF and anti-Caf1 polyclonal antibodies (Fig. 13B, lanes 1; 14C, lane 1). The protein with molecular weight about 18 kDa was also detected with anti-hGMCSF antibodies (Fig. 13B, lane 4). We can presume that this protein is the product of hGMCSF-Caf1 fusion proteolytic degradation.

15 To study influence of the Caf1M chaperone on expression of the *gmcsf-caf1* gene, the *caf1m* gene was inserted into the pACYC-trx, a vector with a low copy number and compatible with pBR322-based plasmids. PACYC-trx was pACYC184 derivative that contained *trx* gene under the control of *tac* promoter. The *Trx* 20 coding region, flanked with *Kpn*I and *Alw*44I sites, was replaced by the DNA fragment coding for the Caf1M chaperone as follows. The *caf1m* gene was amplified by PCR which included Pfu DNA Polymerase, two primers (GTTGTCGGTACCCATTCCGTAAGGAGG) and 5'- GTTAACGTGCACACAGGAACAGC) and the pFS2 plasmid (Galyov EE et al. 25 (1990) FEBS Letters, 277, 230-232). The PCR fragment was treated with *Kpn*I and *Alw*44I and cloned into the *Kpn*I-*Alw*44I large fragment of pACYC-trx. The plasmid obtained was designated as pCaf1M.

JM101 *E. coli* cells were transformed with both pFGMF1 and pCaf1M plasmids 30 followed by analysis of recombinant cell proteins as described above. Co-expression of the chaperone gene from the pCaf1M plasmid led to a marked increase in the amount of the full-length hGMCSF-Caf1 fusion protein in the

periplasm (Fig. 13B, lane 3; 13C, lane 2). The amount of 18 kDa protein, the product of hGMCSF-Caf1 fusion proteolytic degradation, decreased. These results demonstrate that Caf1M promotes the correct folding of hGMCSF-Caf1 and enhances the stability of the fusion protein.

5

#### EXAMPLE 4

##### Co-expression of hIL1ra-Caf1 fusion protein and the chaperone

Based on pFGM13, the pFRA275 plasmid was obtained. In this plasmid the 5'-  
10 terminal part of the mature hIL-1ra coding region containd AlaAspAsp-coding  
sequence instead of the first Arg codon in order to neutralise the N-terminal  
positive charge of hIL1ra. The *lac* promoter controlled the expression of *scaf-*  
*hil1ra* gene in pFRA275.

15 The expression plasmid pFRF275 coding for the *scaf-hil1ra-caf1* gene was  
constructed as shown in Fig. 11. At the 3' terminus of the *hil1ra* gene the *Kpn*2I  
site was introduced by PCR of the pFRA275 plasmid with primers 5'-  
GGAATCCATGGAGGGAAAGAT and 5'-  
ATTATTCCGGACTCGTCCTCCTGAAAGTAG. The amplified fragment was  
20 cut with *Nco*I and *Kpn*2I and ligated with the *Hind*III- *Kpn*2I large fragment from  
pFGMF1 together with the *Hind*III- *Nco*I fragment from pFRA275. The resulting  
plasmid (pFRF275) contained a gene which encodes hybrid precursor consisting of  
the *scaf* signal sequence, *hil1ra* with amino acid changes mentioned above, a  
25 Ser(4GlySer)<sub>3</sub> spacer, and Caf1. Plasmid structure was confirmed by restriction  
analysis and sequencing of amplified regions.

Expression of the *scaf-hil1ra-caf1* gene in JM101 *E. coli* cells was analysed as  
described in EXAMPLE 3. When cells were transformed with pFRF275 alone, the  
30 *hil1ra* fusion protein accumulated mainly in an insoluble form (35 kDa, Fig.12,  
lane 1, 2). However, some part of the fusion protein was translocated to the  
periplasm, where it was subjected to degradation (Fig. 14B, lane 1). When  
pFRF275 and pCaf1M (see EXAMPLE 3) were simultaneously present in

29

recombinant cells, Calf M noticeably diminished the amount of cleavage products (Fig. 14B, lane 2).

5

10

15

20

25

30

**CLAIMS**

1. A bacterial strain simultaneously expressing the fusion protein (signal peptide)-(mature heterologous protein)-(subunit of a bacterial surface structure) and a 5 periplasmic chaperone specific for the subunit for secretion of a recombinant heterologous protein into periplasm of bacteria.
2. A bacterial strain according to Claim 1 additionally expressing outer membrane usher or secretin protein specific for said subunit, for the purpose of secretion of a soluble recombinant heterologous protein on an outer surface of the bacterium or 10 into cultivation medium of the bacterium.
3. A bacterial strain according to Claim 1 or 2, wherein any parts of the fusion protein, periplasmic chaperone and/or outer membrane usher and/or secretin protein are genetically modified.
4. A bacterial strain according to Claim 1, 2 or 3, wherein, instead of the fusion 15 protein, a subunit of a bacterial surface structure with inserted heterologous amino acid sequence(s) is expressed.
5. Use of the bacterial strains according to Claim 1, 2 or 3 for producing heterologous recombinant proteins.
6. Use of the bacterial strains according to Claim 1, 2 or 3, wherein the protein to 20 be expressed is GMCSF, IL-1 $\beta$ , or IL-1 receptor antagonist.
7. Bacterial strain, according to Claim 1, 2 or 3, wherein the microbe is *Escherichia coli*.
8. A microbial expression of recombinant heterologous proteins according to Claim 1, 2 or 3 involving the expression components of Caf1 from *Yersinia pestis*.
- 25 9. A gene-engineered *Escherichia coli* strain containing any of the expression systems of Claim 1, 2 or 3.

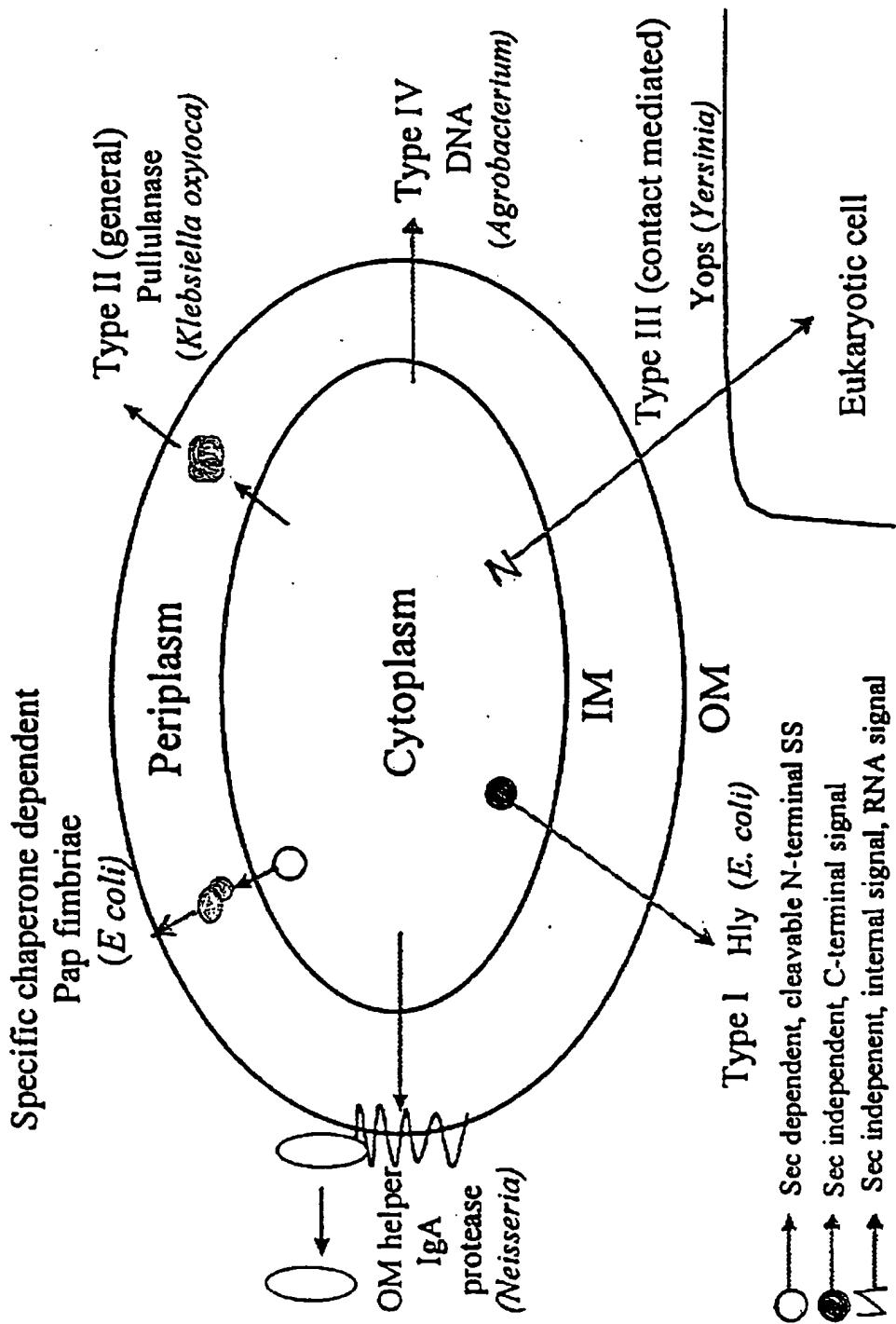


Fig. 1

2/14

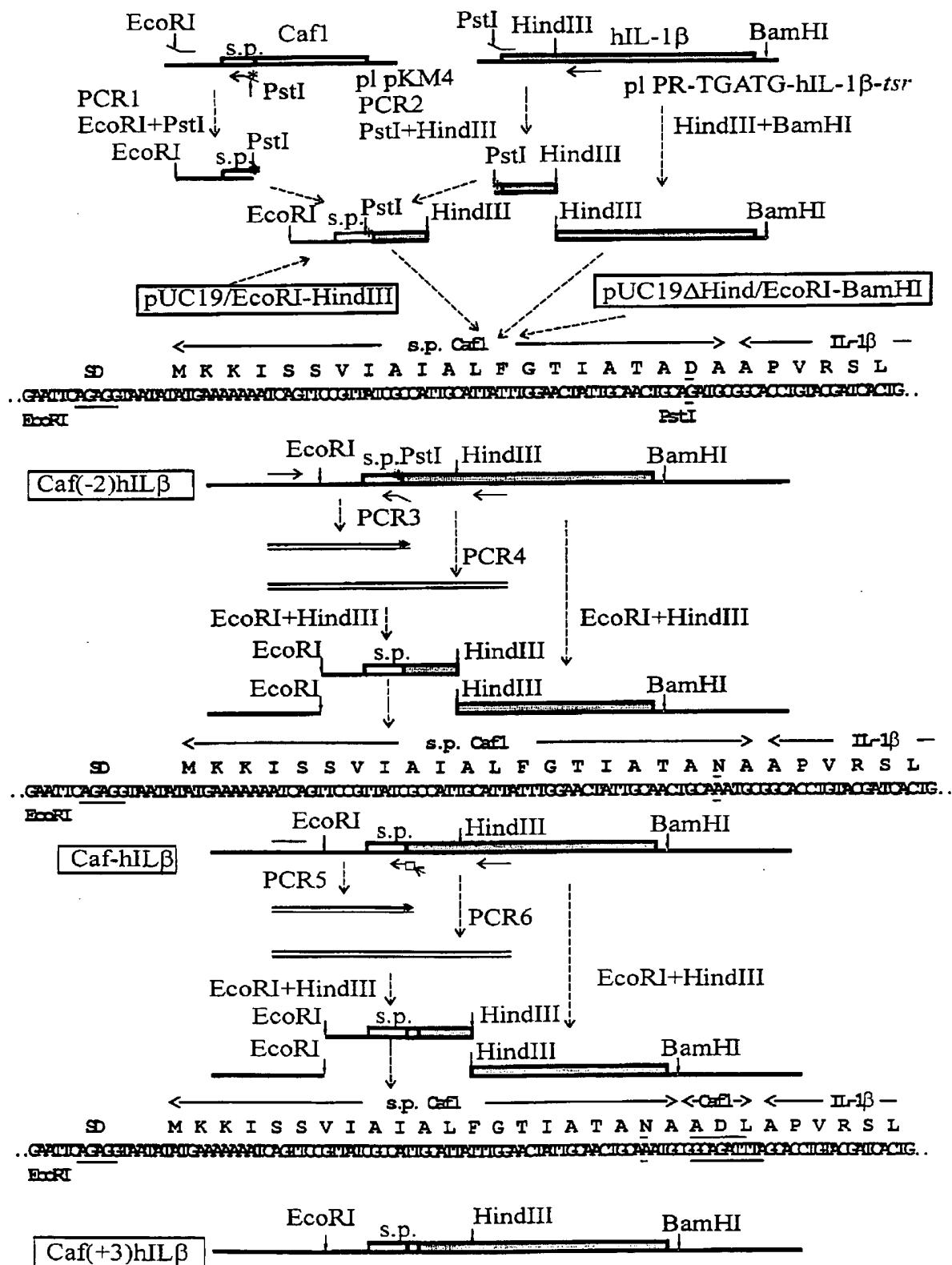
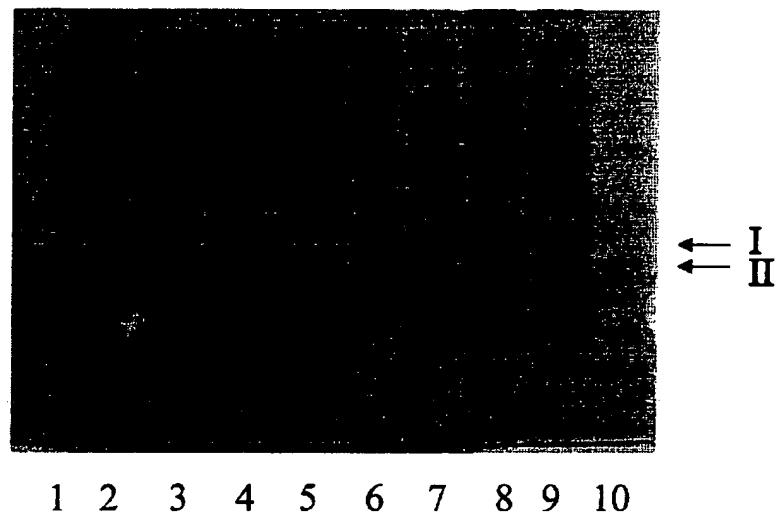


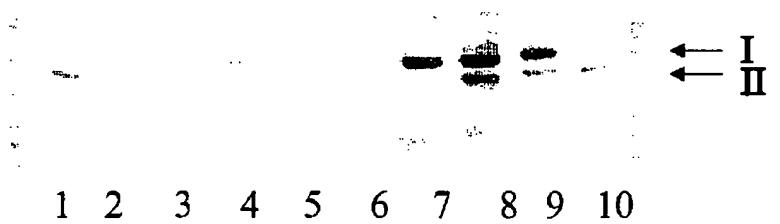
Fig. 2

3/14



1 2 3 4 5 6 7 8 9 10

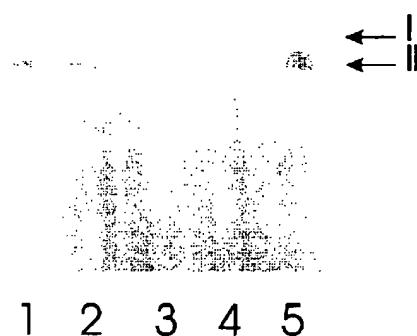
Fig. 3A



1 2 3 4 5 6 7 8 9 10

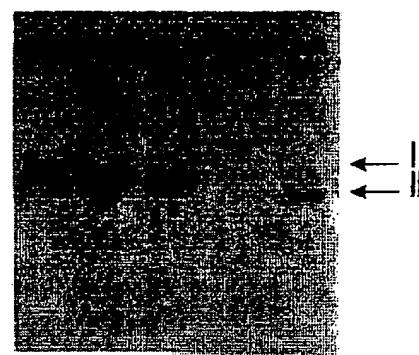
Fig. 3B

4/14



1 2 3 4 5

Fig. 4A



1 2 3 4 5

Fig. 4B

5/14

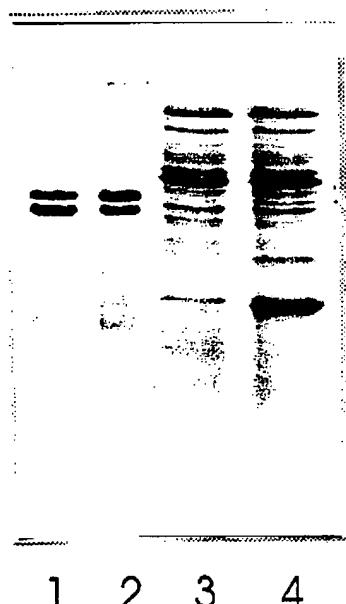
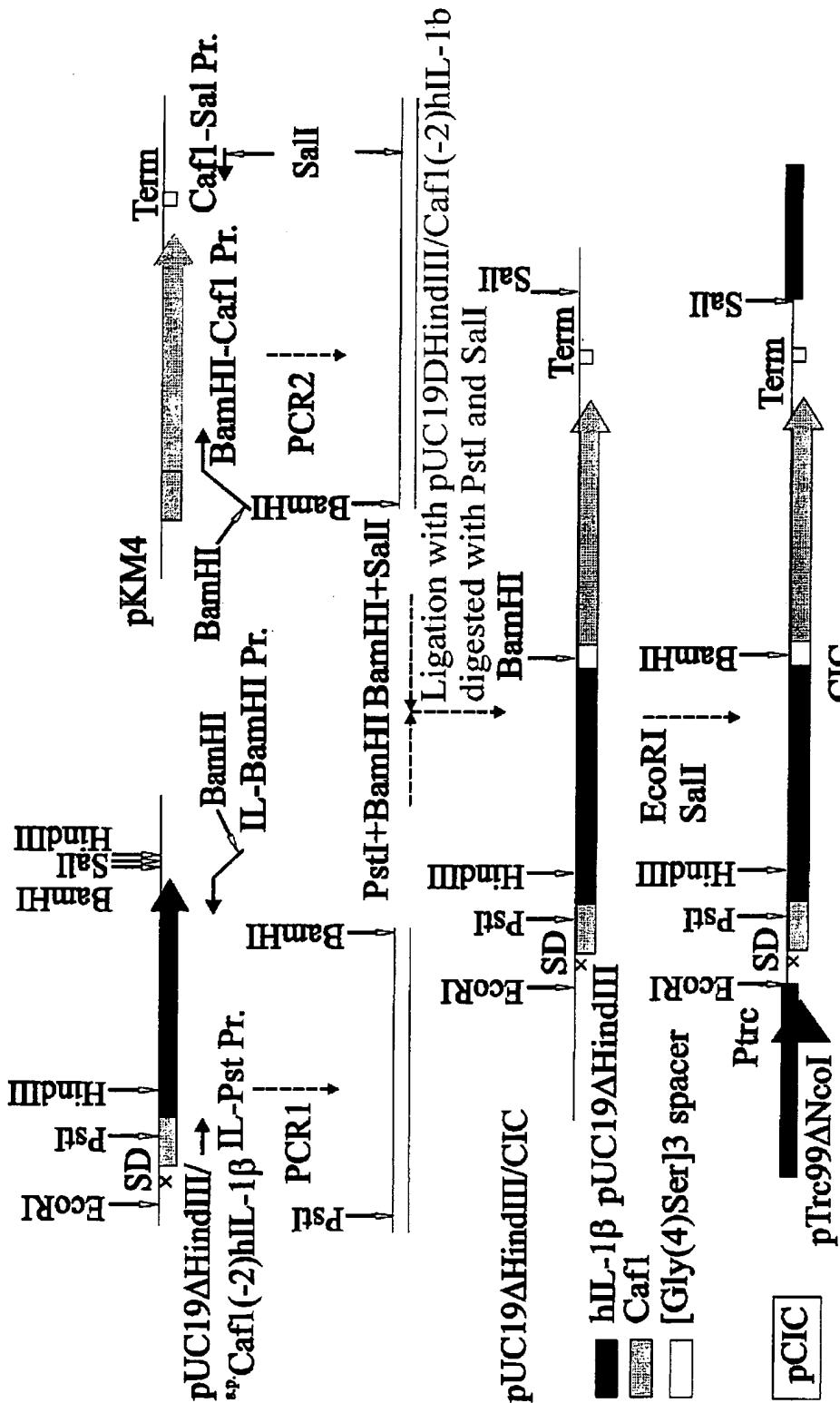


Fig. 5



6  
Fig

7/14

Number	Amino acid
1	Ala
2	Pro
3	Val
4	Arg
5	Ser

Fig. 7

8/14

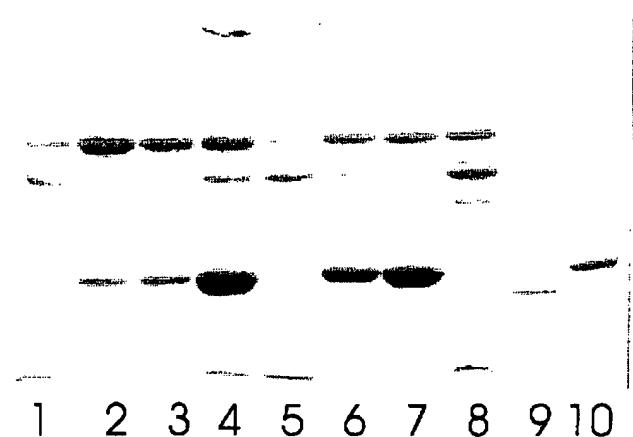


Fig. 8A

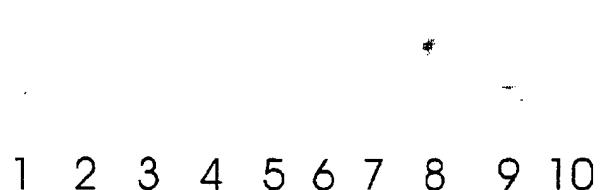


Fig. 8B

9/14

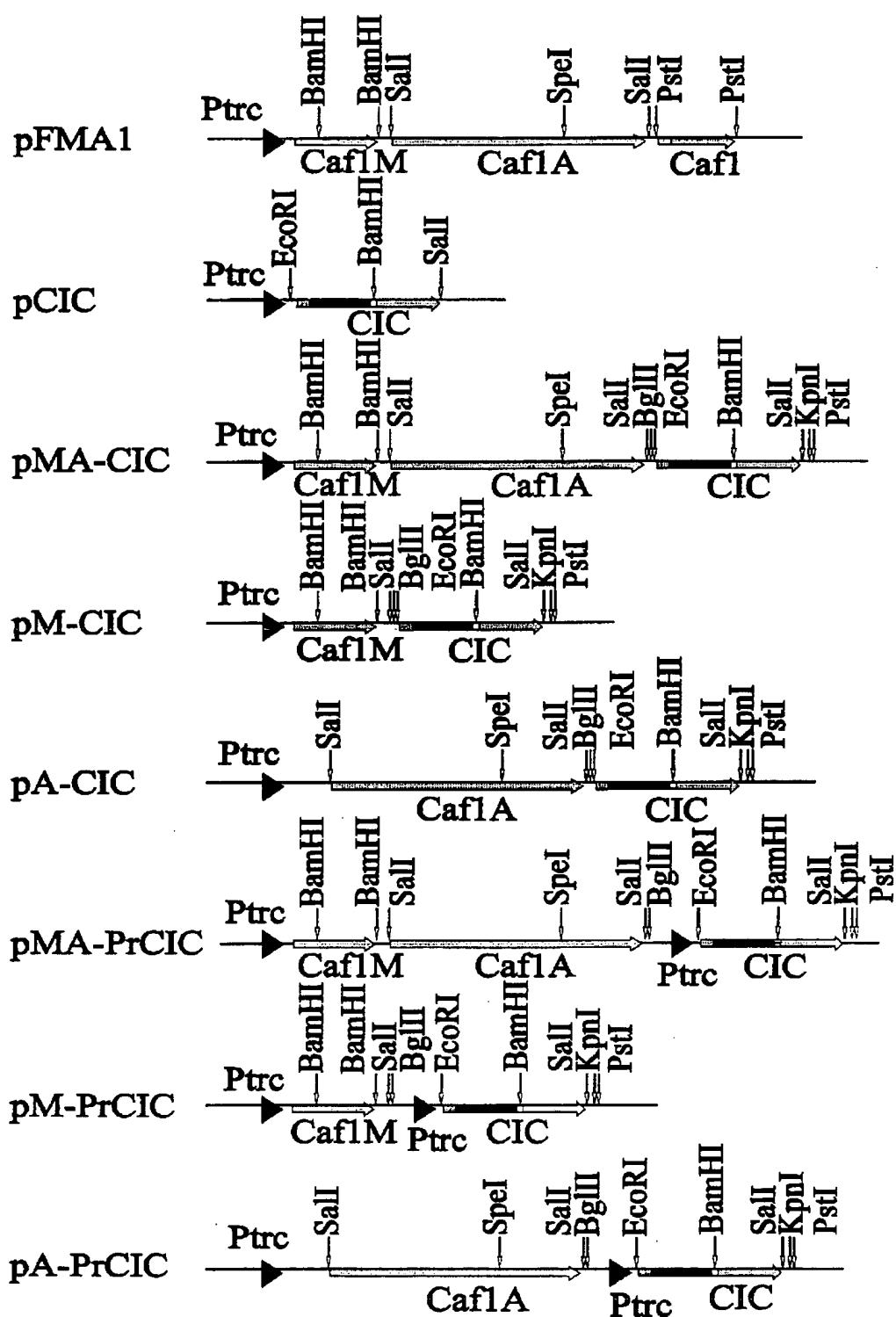


Fig. 9

10/14

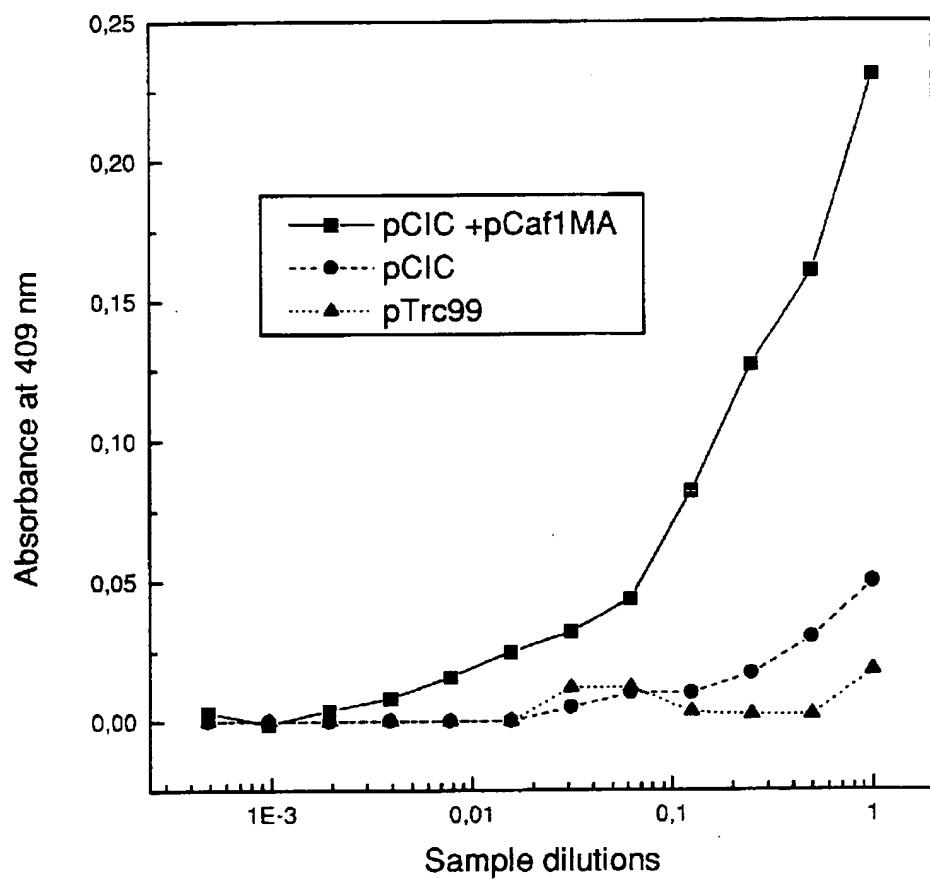


Fig.10

11/14

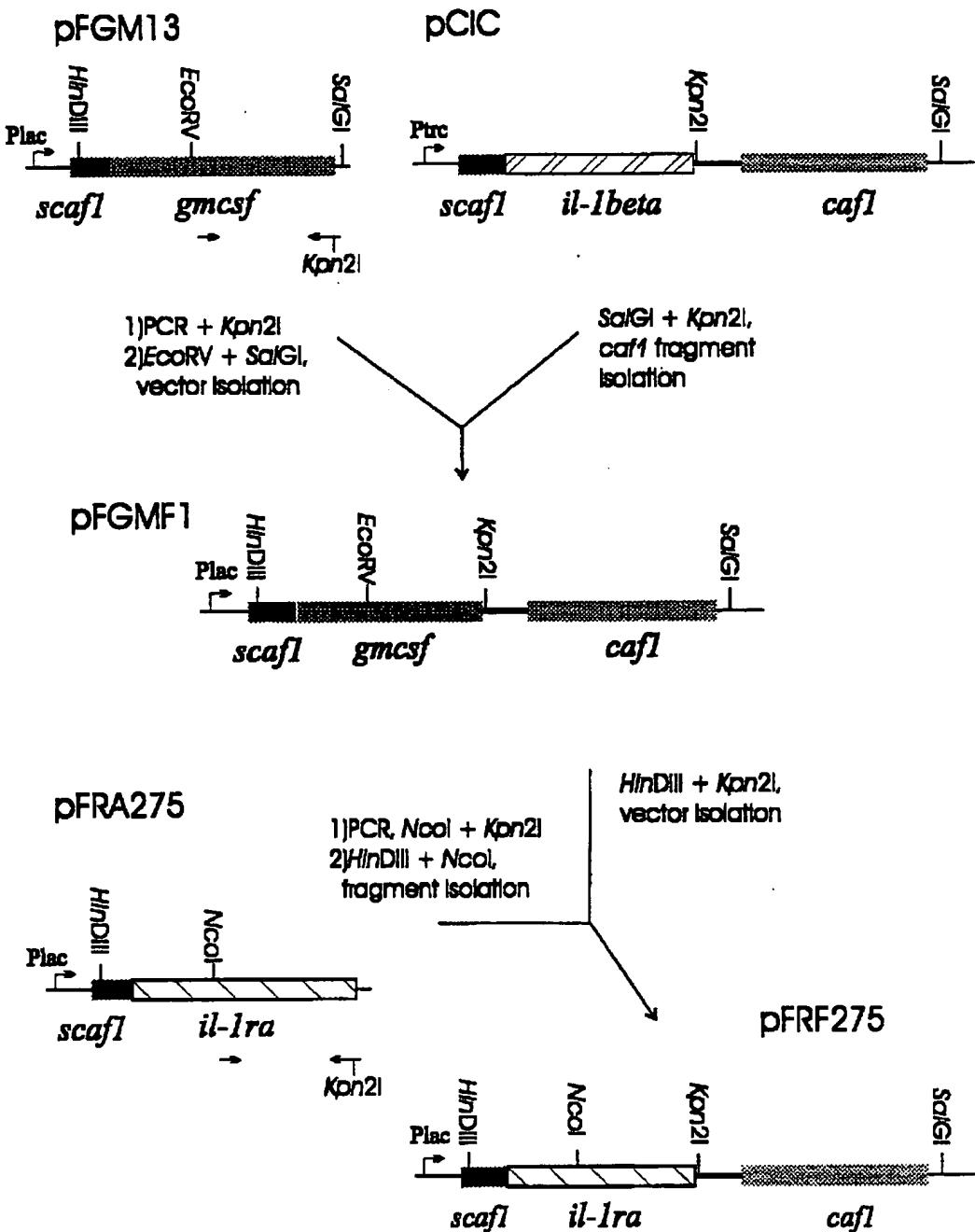


Fig. 11

12/14

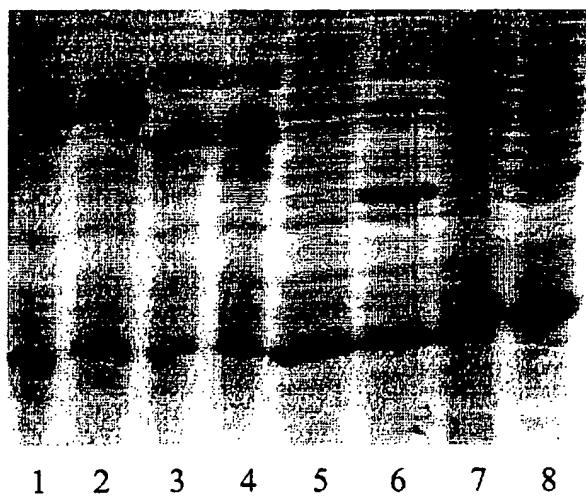


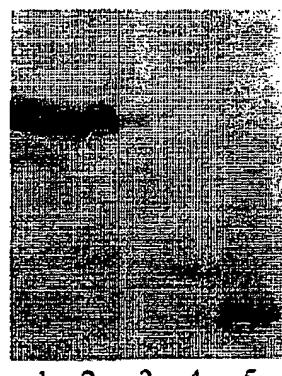
Fig. 12

13/14



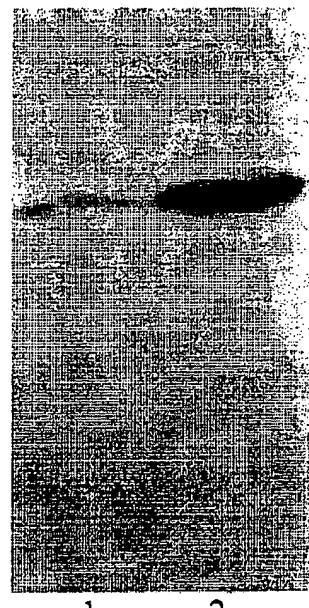
1 2 3

Fig. 13A



1 2 3 4 5

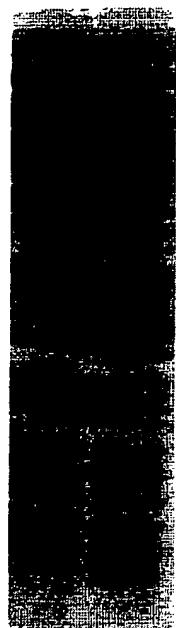
Fig. 13B



1 2

Fig. 13C

14/14



1      2



1      2

Fig. 14A

Fig. 14B

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI 00/0387

## A. CLASSIFICATION OF SUBJECT MATTER

IPC7: C12N 15/70

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	STN International, File MEDLINE, MEDLINEaccession no. 96170663, Petrovskaya L E et al: "Effect of the topography of the signal peptidase site on the effectiveness of secretion of recombinant human granulocyte-macrophage colony-stimulating factor into Escherichia coli periplasm"; & Bioorganicheskaya Khimiia, (1995 Dec) 21 (12) 912-9  --	1-9
A	Biochemical and biophysical research communications, Volume 210, No 2, May 1995, Julián Pérez-Pérez et al, "DNAK/DNAJ supplementation improves the periplasmic production of human granulocyte-colony stimulating factor in escherichia coli", page 524 - page 529, see page 525, lines 2-4  --	1-9

 Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

12 Sept 2000

Date of mailing of the international search report

13-09-2000

Name and mailing address of the ISA/  
Swedish Patent Office  
Box 5055, S-102 42 STOCKHOLM  
Facsimile No. + 46 8 666 02 86

Authorized officer

Yvonne Siösteen/EÖ  
Telephone No. + 46 8 782 25 00

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI 000387

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0885967 A2 (HSP RESEARCH INSTITUTE, INC.), 23 December 1998 (23.12.98), see abstract	1-9
	--	
A	EP 0774512 A2 (IMANAKA, TADAYUKI), 21 May 1997 (21.05.97), see page 4, lines 2-3	1-9
	--	
A	WO 9408012 A1 (RESEARCH CORPORATION TECHNOLOGIES, INC.), 14 April 1994 (14.04.94)	1-9
	--	
A	WO 9614422 A1 (GENENTECH, INC.), 17 May 1996 (17.05.96)	1-9
	-----	

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No.

PCT/FI 00/0387

28/06/00

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP 0885967 A2	23/12/98	CA JP	2235468 A 11009274 A	20/12/98 19/01/99
EP 0774512 A2	21/05/97	JP	9173078 A	08/07/97
WO 9408012 A1	14/04/94	EP US	0663010 A 5773245 A	19/07/95 30/06/98
WO 9614422 A1	17/05/96	CA EP IL JP US US	2203373 A 0786009 A 115834 D 10508203 T 5639635 A 5789199 A	17/05/96 30/07/97 00/00/00 18/08/98 17/06/97 04/08/98

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.  
As rescanning these documents will not correct the image  
problems checked, please do not report these problems to  
the IFW Image Problem Mailbox.**